DOCUMENT RESUME

ED 076 701 TM 002 710

AUTHOR Villaneuva, Lourdes S.; Dunn-Rankin, Peter
TITLE A Comparison of Ranking and Rating Methods by Multidimensional Matching.
PUB DATE 73
NOTE 7p.; Paper presented at annual meeting of American Educational Research Association (New Orleans, Louisiana, February 25-March 1, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS College Students; *Comparative Analysis; *Evaluation Methods; Higher Education; *Multidimensional Scaling; Orthogonal Rotation; Rating Scales; *Sorting Procedures; Technical Reports

ABSTRACT

Ranking and rating methods of making direct judgments of pairwise similarity are compared. Three dimensional configurations of seven objects are derived under both judgment methods across two equivalent groups of 59 university students. Using Cliff's orthogonal rotation method of multidimensional matching, comparisons are made between methods (Ranking vs. Rating for Groups 1 and 2) and between groups (Group 1 vs. Group 2 for Ranking and Rating). The comparisons between methods reveal that ranking and rating have highly congruent configurations for both groups of subjects. While the comparisons between groups in both methods are highly congruent, ranking produced greater congruency for the third dimension. (Author)

A COMPARISON OF RANKING AND RATING METHODS BY MULTIDIMENSIONAL MATCHING

Ma. Lourdes S. Villanueva
Peter Dunn-Rankin
Department of Educational Psychology
University of Hawaii

While there exist a variety of procedures for collecting similarities data which are analyzed by nonmetric multidimensional scaling techniques, there also exists the problem of finding out the extent to which these procedures are comparable so that they may be efficiently employed. Green and Carmone (1970) state that future research in the analysis of similarities data calls for the determination of whether or not alternative data gathering procedures, as presently applied, yield invariant similarity configurations.

Since the most widely used methods for obtaining direct judgments of similarity are ranking and rating, it appears worthwhile to investigate whether or not the separate multidimensional scaling of the ranking and rating responses to the same pairs of psychological objects results in congruent configurations.

The rating task usually involves assigning estimates of similarity for each pair of objects on a scale of some stated range, while ranking calls for the stepwise ordering of pairs of objects in terms of increasing or decreasing similarity. Past experience indicates that while ranking maximizes the transitivity of pairs on the individual response level, it is, however, the more demanding of the two tasks and the difficulty it presents increases dramatically with the increase in the number of psychological objects, thereby allowing for an increase in the occurrence of undesirable response sets. Rating, on the other hand, seems to be less demanding, requiring much shorter time for completion. The problem, then, is to determine whether rating can adequately substitute for ranking.

Method

Judgment of similarity were collected for all pairs of seven campus "places" using both rating and ranking methods. The similarity configurations derived from both methods were then matched.

Subjects

Two classes of 59 graduate students at the University of Hawaii served as Ss. The groups had similar age and sex composition and were motivated to carry through the tasks by their course instructors. Class A (n = 28) rated first and ranked second while Class B (n = 31) ranked first and rated second.

Stimuli

The seven campus "places" with respect to which judgment of similarity were sought were: classroom, dormitory, library, cafeteria, gymnasium, theater, and laboratory. The sequence of pairs, as they appeared in the rating sheet and the ranking pile, followed the optimum presentation order recommended by Ross (1934) and was the same for both tasks and for all Ss.

Procedure

For the ranking task, each S was instructed to sort the pairs into two piles consisting of (1) similar and (2) dissimilar pairs of objects. This initial sorting procedure has been found effective in simplifying the ranking task (Green & Carmone, 1970). The S was instructed to rank order the pairs in both piles according to the similarity between the two paired "places" on each card. The position of any one pair in the pile from the most similar to the most dissimilar determined its rank.

For the rating task, each S was instructed to assign any number between zero to 100 to each pair according to how similar the two places appeared. The most similar pair would have the highest rating while the most dissimilar would have the lowest rating. The 100-point rating scale was chosen because of its familiarity and relatively wide range.

Analysis

The mean rating and mean ranking of each pair of "places" were obtained from each of the four sets of data: (1) Class A - Rate First, (2) Class A - Rank Second, (3) Class B - Rank First, and (4) Class B - Rate Second. The means constituted the ordinal distance data input to Kruskal's (1964a, b) multidimensional scaling program called "M-D-SCALE." The number of dimensions which produced the best fitting configuration was determined according to the amount of stress present in the solutions. Stress is a numerical value which denotes the degree of departure of the observed similarity from the true similarity among objects taken two at a time (Kruskal, 1964a, b).

To test the equivalence of the two methods, the coordinates of the "places" in the final best-fitting configurations obtained by ranking were matched to those obtained by rating. To test the equivalence between Class A and Class B so that the presentation order effect of the tasks can be determined, the coordinates of the "places" in the final best-fitting configurations of Class A were compared to those of Class B. Cliff's (1966) least squares method for orthogonal rotation to congruence provided an objective method for multidimensional matching. The degree to which the two configurations are mutually fitted is expressed by the statistic known as coefficient of congruence[1] (Harman, 1960).

Results and Discussion

The best fitting solutions obtained in 1, 2, 3 and 4 dimensions for each of the four sets of data are shown in Table 1. Since the amount of stress did not decrease considerably after three dimensions, then, the three-dimension solutions which denoted between "good" to "excellent" fit (Kruskal, 1964a) were found to be adequate representations of the data. While the ranking solutions had lesser stress than the rating solutions, the differences were not significant.

Insert Table 1 about here.

[1] Nesselroade and Baltes (1970) discuss the difficulty in using the coefficient of congruence as an index of dimensional or factorial similarity.

Dimension I is a VIGOROUS ACTIVITY dimension which polarized gymnasium against library and dormitory. Dimension II isolated theater. Dimension III is a STUDY ACTIVITY dimension which grouped classroom, library and laboratory against cafeteria, dormitory and gymnasium. These dimensions were similarly structured for the four sets of data.

Multidimensional Matching

Between Classes. Class A and Class B had highly congruent configurations in both rating and ranking methods. Table 2 shows that the Class A - Class B coefficients of congruence (CC) in rating for Dimensions I, II and III were .97, .94 and .58, respectively; while the Class A - Class B CCs in ranking were .95, .87 and .89, respectively. This indicates that Class A and Class B appear to have very alike judgments on the similarity of the campus "places," ~~particularly in Dimensions I and II. This lends~~ support to the attitudinal equivalence of the classes, in addition to their pre-experimental equivalence in sex, age and education level composition.

Insert Table 2 about here.

Figure 1 shows the two-space configuration of Class A superimposed on Class B in rating and ranking. The squares denote plots of objects obtained by ranking while circles denote those obtained by rating. Clear squares and circles indicate that the task was performed first while shaded ones indicate that the task was completed second.

Insert Figure 1 about here.

Between Methods. The configurations derived by the rating method were highly congruent to those derived by the ranking method for both Class A and Class B. Table 2 shows that the rating versus ranking coefficients of congruence (CC) in Class A for Dimensions I and II were .98 and .92, respectively; while in Class B, they were .97 and .96, respectively. Since these CCs were close to unity, they indicate that there were no appreciable systematic difference between the ranking and rating configurations for Dimensions I and II. In Dimension III, however, Class A (which rated first) had a higher CC than Class B (which ranked first). Since Class A and Class B only had moderate mutual fit in Dimension III, it can not be stated conclusively that rating objects first tended to produce more congruent configurations.

Conclusions

The dimensions underlying the three-dimensional solutions are similar between classes and between methods in that Dimension I is a VIGOROUS ACTIVITY dimension, Dimension II is a THEATER dimension, and Dimension III is a STUDY ACTIVITY dimension.

Rating appears to be an adequate alternative method of judging pairwise similarity. With familiar objects like campus "places" and sophisticated judges such as the graduate students who served as subjects in this study, rating produces configurations which highly match the ones produced by ranking.

Rating objects first seems to make rating and ranking configurations more congruent while the reverse is not observed.

Educational Implications

With sophisticated judges and familiar objects, rating can be substituted for ranking with both methods producing similar configurations under multidimensional scaling. Under similar conditions, the use of the rating method will economize the time and effort needed for tasks requiring judgment of similarity. Since the tediousness of the ranking task increases the occurrence of undesirable response sets and random responding particularly with young Ss, it seems worthwhile to investigate whether rating can substitute for ranking among children.

References

Cliff, Norman. "Orthogonal rotation to congruence", Psychometrika, 1966, v. 31, n. 1, p. 33.

Harman, Harry H. Modern factor analysis. Chicago: University of Chicago Press, 1960, 471 pp.

Green, Paul E. and Frank J. Carmone. Multidimensional Scaling. Boston: Allyn and Bacon, 1970, 203 pp.

Kruskal, J. B. "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis", Psychometrika, 1964, v. 29, n. 1, p1.

Kruskal, J. B. "Nonmetric multidimensional scaling: a numerical method", Psychometrika, 1964, v. 29, n. 2, p. 115.

Nesselroade, John R. and Paul B. Baltes. "On a dilemma of comparative factor analysis: a study of factor matching based on random data", Educational and Psychological Measurement, 1970, v. 30, p. 935.

Ross, Robert T. "Optimum orders for the presentation of pairs in the method of paired comparisons", Journal of Educational Psychology, 1934, v. 25, p. 375.

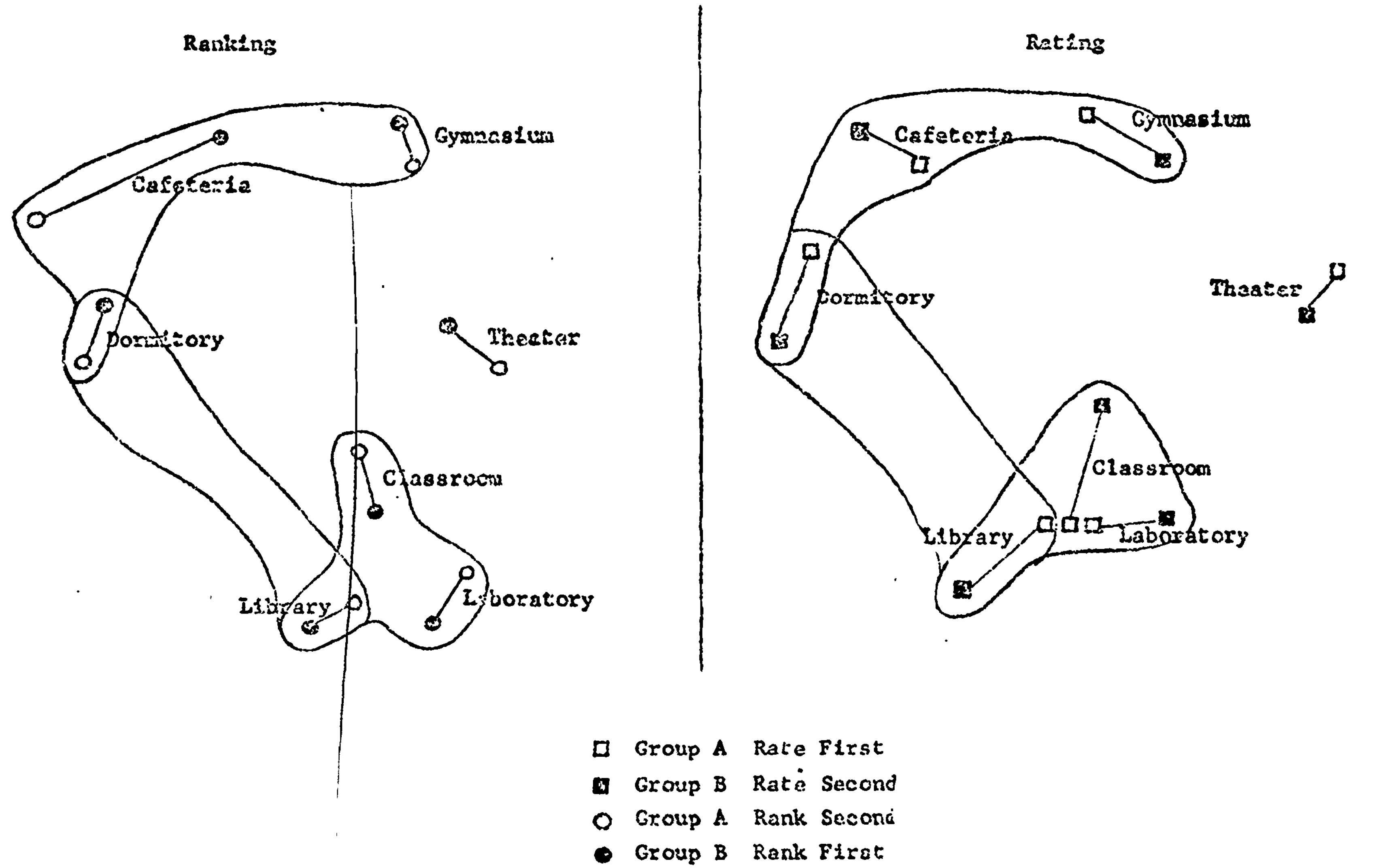

FIG. 1 Superimposed Group A and Group B Two-Space Configurations.

Table 1

Stress Across Four Dimensions

Data Groups	Number of Dimensions			
	1	2	3	4
Class A - Rate First	.678	.281	.025**	.030*
Class A - Rank Second	.324	.142	.008**	.006**
Class B - Rank First	.386	.166	.009**	.007**
Class B - Rate Second	.515	.154	.029*	.037*

** Goodness of fit $\geq$ "excellent"

* Goodness of fit $\geq$ "good"

Table 2

Coefficient of Congruency

Data Groups	Dimensions		
	I	II	III
Between Classes:			
Class A - Rate First vs. Class B - Rate Second	.97	.94	.58
Class B - Rank First vs. Class A - Rank Second	.95	.93	.89
Between Methods:			
Class A - Rate First vs. Class A - Rank Second	.98	.92	.98
Class B - Rank First vs. Class B - Rate Second	.97	.96	.61